LIFTING-BASED DIRECTIONAL DCT-LIKE TRANSFORM FOR IMAGE CODING

Hao Xu[1], *Jizheng Xu*[2], *Feng Wu*[2]

[1]Dept. Electronical Engineering & Information Science, Univ. of Science and Technology of China, Hefei, China
[2]Internet Media Group, Microsoft Research Asia, Beijing, China

ABSTRACT

Traditional 2D DCT implemented by separable 1D transform in horizontal and vertical directions does not take image orientation features in a local window into account. To improve it, we propose to introduce directional primary operations to the lifting-based DCT and thereby derive a new directional DCT-like transform. A JPEG-wise image coding scheme is also proposed to evaluate the performance of the proposed directional DCT-like transform. In this scheme, each 8x8 block can be transformed along different direction in order to better adapt to local orientation properties of image. The experimental results show that the performance of the proposed directional DCT can dramatically outperform the conventional DCT up to 2dB even without modifying entropy coding.

Index Terms—directional transform, DCT.

1. INTRODUCTION

Most of image compression schemes are constructed on the identical architecture of 2D transform following with entropy coding, e.g. JPEG uses 2D DCT and VLC, and JPEG2000 uses 2D Wavelet transform and EBCOT. Either 2D DCT or 2D wavelet transform is implemented by separable 1D transform in horizontal and vertical directions. A serious drawback of these transforms is that they are ill suited to approximate image features with arbitrary orientation that is neither vertical nor horizontal. In these cases, they result in large-magnitude high-frequency coefficients. At low bit rates, the quantization noise from these coefficients is clearly visible, in particular causing annoying Gibbs artifacts at image edges with arbitrary directions.

How to incorporate directional information into transform is a challenge in image and video coding. Some work has been reported on wavelet and subband transform. The lifting structure developed by W. Sweldens provides a good way to incorporate directional information into wavelet transform, in which each FIR (Finite Impulse Response) wavelet filter can be factorized into several lifting stages [1]. The proposed ADL (adaptive directional lifting) in [2][3] can outperform JPEG 2000 up to 2dB on images with rich orientation features by first introducing directional data coding and sub-pixel spatial prediction. An alternative approach of exploiting directional information is proposed in [8], where the signal is rotated using circular shifts to fit the transform instead of adapting the transform to signal.

To the best of authors' knowledge, the paper from Zeng et al. is the first effort about how to incorporate directional information into DCT [4]. Their directional DCT is motivated by SA-DCT (shape-adaptive DCT). For an N-by-N block, the first 1D DCT is performed along the direction selected, where each column is of a different length. Then DCT coefficients are reorganized according to their frequencies from DC to AC. Coefficients at the same frequency are viewed as a row and transformed by the second 1D DCT, where each row is also of a different length. Since the directional transform is constrained within one block, direction selection is relatively easy in Zeng's scheme. But this also potentially hurt the performance of directional DCT because the correlation among neighboring blocks with similar direction is not exploited. In addition, directions at fractional pixels are not considered yet in [4].

Inspired by the lifting-based transform [1] and our previous work on ADL [2][3], a lifting-based directional DCT-like transform is proposed in this paper, which can be performed along arbitrary direction in theory. The lifting scheme factorizes 1D DCT into a series of so-called primary operations. We propose the corresponding directional form of each primary operation because introducing directional information in primary operation is much easier than in a whole transform. All the directional primary operations construct the proposed directional DCT-like transform. The perfect reconstruction property is guaranteed by the lifting scheme when transform coefficients are not quantized at all. Since the 8-point DCT is extensively used in image coding, only the 8-point directional transform is taken as an example to discuss in this paper. The proposed ideas can be easily applied to a DCT of arbitrary size.

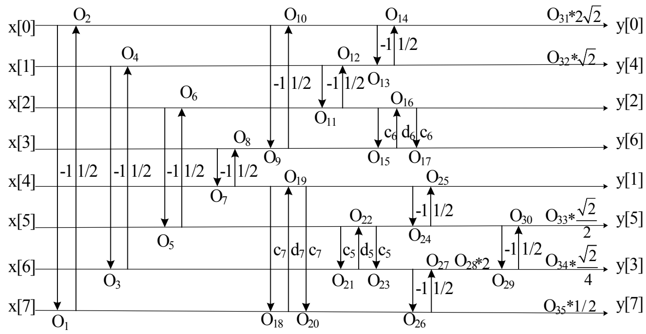
The rest of this paper is organized as follows. The lifting-based directional DCT-like transform is introduced in Section 2. In Section 3, an image coding scheme adopting our directional transform is proposed. The experimental results are presented in Section 4. Finally, the paper is concluded in Section 5.

2. LIFTING-BASED DIRECTIONAL DCT-LIKE TRANSFORM

There are many approaches reported to factorize any N-point 1D DCT into a series of operations through plane rotation and butterflies for the fast implementation [5]. They

[1] This work has been done while the author is with Microsoft Research Asia.

can be further converted to lifting structures as reported in [6]. Fig 1 is one of the lifting structures of 8-point DCT, which is chosen as the basis to develop our directional DCT-like transform in this paper.



$$c_n = \frac{\cos(\frac{n\pi}{16})-1}{\sin(\frac{n\pi}{16})} \quad d_n = \sin(\tfrac{n\pi}{16})$$

Fig 1: Factorization of 8-point DCT into primary operations. O_i ($i = 1, 2, ..., 35$) is primary operation.

In the lifting structure depicted in Fig 1, an 8-point DCT is factorized into 35 primary operations, which can be formalized as

$$\mathbf{Y} = DCT(\mathbf{X}) = O_{35} \circ O_{34} \circ ... \circ O_2 \circ O_1(\mathbf{X}). \quad (1)$$

Except for O_{28} and O_{31}~O_{35}, the primary operations with a parameter α can be expressed as

$$O(x[n_i], x[n_j], \alpha) = \{x[n_i] \leftarrow x[n_i] + \alpha x[n_j]\}. \quad (2)$$

Here, we call $x[n_i]$ and $x[n_j]$ as the target and source of a primary operation respectively for the convenience of later discussion.

Each of these primary operations can be easily extended into directional version. As depicted in Fig 2, if the source of a primary operation locates at fractional pixel, its value is interpolated from neighboring integer pixels in a row. If the target of a primary operation locates at fractional pixel, the resulting value is only an intermediate value. Similar to the energy distributed update method proposed by [7], the intermediate value is immediately distributed to its neighboring integer pixels according to their weighting factors of interpolation. Therefore, primary operation of 1D directional transform is actually acted on 2D signal. The primary operation along a given direction θ is expressed as

$$\tilde{O}(X[n_i,m], F(X,n_j,\theta), \alpha) = \{X[n_i,m] \leftarrow X[n_i,m] + \alpha F(X,n_j,\theta)\} \quad (3)$$

$$F(X,n_j,\theta) = \begin{cases} X[n_j, m+(n_i-n_j)\tan(\theta)] & (n_i-n_j)\tan(\theta) \text{ is integer} \\ \sum_{k=-T+1}^{T} w_k X[n_j, m+\lfloor (n_i-n_j)\tan(\theta) \rfloor + k] & \text{otherwise} \end{cases}$$

where X denotes a 2D array of pixels and $X[n,m]$ denotes a pixel located at row n and column m. $\lfloor . \rfloor$ is the floor function and w_k's are the parameters of a $2T$-tap interpolation filter. The illustrations of O_5 and $\tilde{O}_5$ are shown in Fig 2 (a) and Fig 2 (b) respectively, where α=-1 and θ=arctan(1/2).

Because only one pixel is involved in the primary operations denoted as O_{28} and O_{31}~O_{35}, their directional versions are independent of directional parameter θ, that is

$$\tilde{O}(X[n,m], \alpha) = O(X[n,m], \alpha) = \{X[n,m] \leftarrow \alpha X[n,m]\}. \quad (4)$$

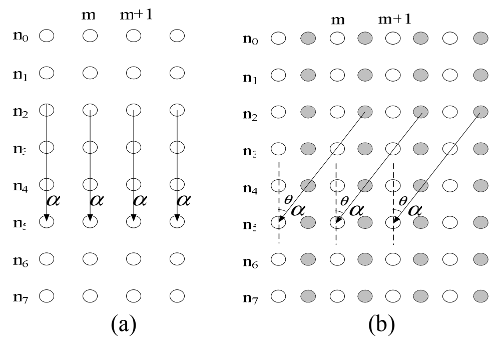


Fig 2: Exemplified primary operations of (a) non-directional and (b) directional, where white circles denote integer pixels and gray circles half pixels.

It is not difficult to derive a new directional transform by reassembling the directional primary operations according to the order and structure depicted in Fig 1. The directional transform can be generally described as follows

$$\mathbf{Y} = DCT_{dir8}(\mathbf{X}, \theta) = \tilde{O}_{35} \circ \tilde{O}_{34} \circ ... \circ \tilde{O}_2 \circ \tilde{O}_1(\mathbf{X}, \theta). \quad (5)$$

DCT_{dir8} indicates 1D 8-point directional DCT-like transform and $\mathbf{X}$ is a pixel matrix of size 8×M. The directional parameter of every primary operation is θ. $\tilde{O}_i$ is formed by the operations defined in equations (3) and (4) where $m = 1...M$. Accordingly, $\mathbf{Y}$ is a coefficient matrix of size 8×M. The corresponding inverse 1D directional DCT-like transform is given as

$$\mathbf{X} = DCT_{dir8}^{-1}(\mathbf{Y}, \theta) = \tilde{O}_1^{-1} \circ \tilde{O}_2^{-1} \circ ... \circ \tilde{O}_{34}^{-1} \circ \tilde{O}_{35}^{-1}(\mathbf{Y}, \theta) \quad (6)$$

where $\begin{cases} \tilde{O}^{-1}(X[n_i,m], F(X,n_j,\theta), \alpha) = \tilde{O}(X[n_i,m], F(X,n_j,\theta), -\alpha) \\ \tilde{O}^{-1}(X[n_i,m], \alpha) = \tilde{O}(X[n_i,m], \frac{1}{\alpha}) \end{cases}$.

The proposed directional DCT-like transform can guarantee the perfect reconstruction property if transform coefficients are not quantized at all, namely,

$$\mathbf{X} = DCT_{dir8}^{-1}(DCT_{dir8}(\mathbf{X}, \theta), \theta), \quad (7)$$

because the primary operations defined in equations (3) and (4) are all perfectly reversible.

3. CODING SCHEME WITH PROPOSED DIRECTIONAL TRANSFORM

A JPEG-wise coding scheme is proposed to evaluate the performance of our directional DCT-like transform, as depicted in Fig 3. The modules that do not relate to transform are the same as those in JPEG, such as quantization and entropy coding.

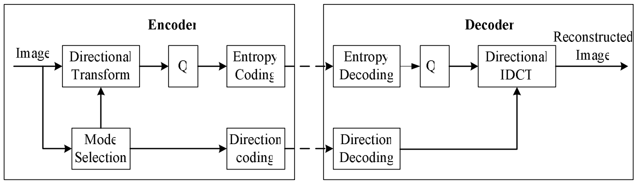


Fig 3: The coding scheme using directional DCT-like transform.

In the scheme, approximately-optimal transform directions for blocks are extracted as side information in "mode selection" according to rate-distortion criterion. The side information is losslessly coded by predicted VLC method and transmitted to decoder.

The directional transform employed in our scheme is a 2D extension of the 1D directional DCT-like transform introduced in the previous section. Actually, after the 1D directional transform, generated coefficients do not own orientation property again. They are organized in different rows, saying DC, AC_1, … , and AC_7, from low frequency to high frequency. Then, the second 1D transform is a normal one and performed on each row. For simplification, the transforms are vertically aligned, so that they are well filled in the 8x8 blocks.

The available modes are defined in Fig 4 as the set $\{-4,...,0,...,4\}$, where integer pixels are marked by "○", half pixels by "+" and quarter pixels by "×".

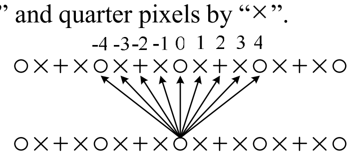


Fig 4: Pre-defined available direction modes

3.1. Direction transition on block boundary

To adapt to local orientation properties of image, each block has its own transform direction. If the neighboring blocks have different directions, the direction transition will take place on their boundary. For the convenience of discussion, we draw the "direction threads" on blocks to denote their directions as shown in Fig 5 as dashed lines. Actually, the direction threads function as a hint to associate the target and source of any directional primary operation.

The transition can be categorized into three cases according to our pre-defined modes. The first case is that two blocks both have positive modes and left block's direction threads can reach its rightmost column of integer pixels, as exemplified in Fig 5 (a) (symmetrical case is also included). In this case, these two groups of direction threads match well so that the transition goes smoothly.

The second case is similar to the first case, except that some of left block's direction threads do not reach an integer pixel on the last column before going into the boundary area (the area between two vertical dash-dot lines in Fig 5). Accordingly, we change their directions at the boundary area so that each of them links to one integer pixel on the first column of right block, as exemplified in Fig 5 (b).

In the third case, these two blocks have the modes with opposite sign, or one of direction modes in these two blocks is zero. The two blocks show little directional correlation in this case. Therefore, we turn the direction threads to the vertical direction along with the last column as exemplified in Fig 5 (c).

However, there exists multi-to-one mapping during the transition of the 2nd and 3rd cases. In the primary operation of prediction (downward arrows in Fig 1), the unique source can be found along direction thread, no matter whether multi-to-one mapping exists or not. But in the primary operation of update, multiple sources may be used to update a target according to the direction threads. Fortunately, this problem has been solved by the EDU (energy distributed update) proposed in [7]. We adopt the same approach in this paper.

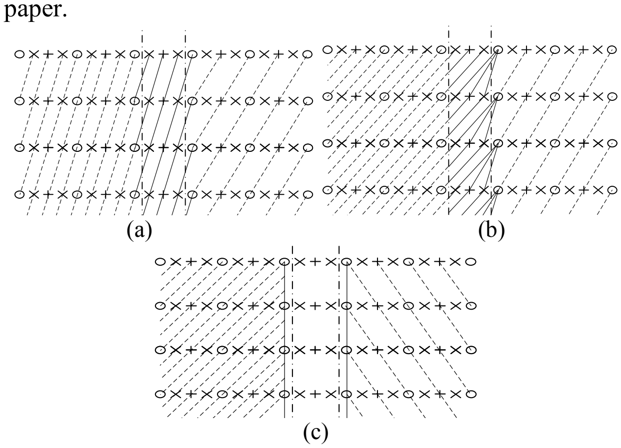


Fig 5: Exemplified direction transitions

3.2. Mode decision

Mode decision is aiming to select optimal transform directions for blocks so that the overall coding performance can be maximized. In our scheme, we adopt a mode decision approach basing on dynamic programming algorithm. Due to the space limit, we are not going to introduce it in detail here. More details on this are available in [9].

4. EXPERIMENTAL RESULTS

In this section, we provide the experimental results to compare the performance of our directional DCT-like transform with conventional DCT. The testing images include Lena(512x512), Barbara(512x512), Peppers(512x512), Man(512x512), Photographer(256x256) as well as the first frame of Foreman(352x288). There are 9 available modes as shown in Fig 4 and the 8-tap Sinc filter is used as the interpolation filter.

The PSNR/bpp curves are depicted in Fig 6, while the coding quality is set from 20 to 90 with incremental step 10. The coding gain can be up to 2.0dB for Barbara. On the testing set the gain of the proposed directional DCT-like transform ranges from 0.2dB to 2.0dB, depending on the presence or lack of orientation features in the image. Some visual quality comparisons at same bit-rate are given in Fig 7.

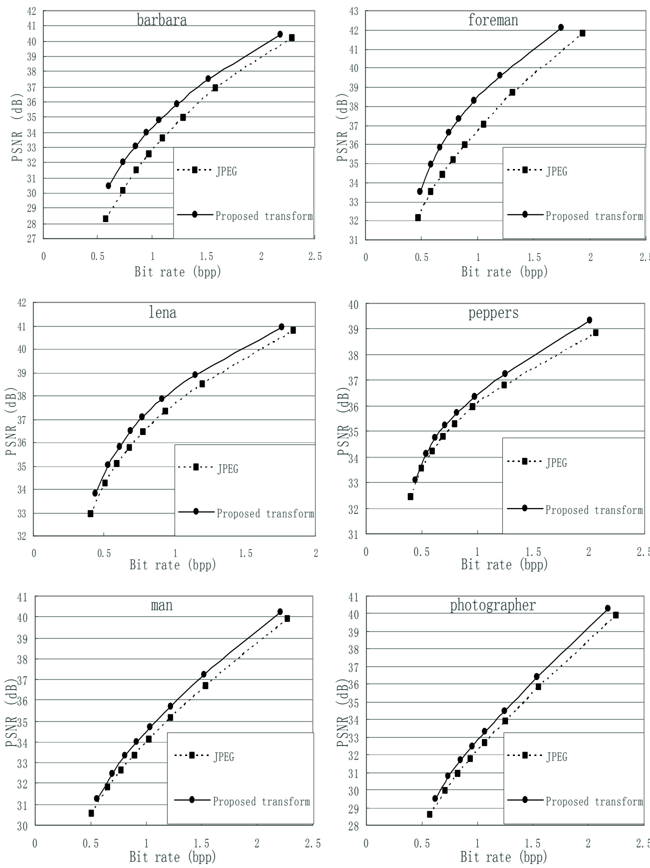


Fig 6: PSNR/bpp curves

5. CONCLUSION

In this paper, we develop a directional DCT-like transform basing on the lifting structure of DCT. The transform can be performed along arbitrary direction in theory and still has the perfect reconstruction property. At the same time, a JPEG-wise coding scheme adopting such directional transform is proposed, in which each 8x8 block is transformed along variable direction in order to better adapt to local orientation properties of image. Experimental results show that the proposed directional DCT-like transform significantly outperforms conventional DCT in terms of PSNR/bpp and visual quality. For more details and analysis, the reader can refer to [9].

Fig 7: Visual quality comparison. (left column: JPEG; right column: our scheme)

6. REFERENCES

[1] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," Technical Report 1994:7, Industrial Mathematics Initiative, Department of Mathematics, University of South Carolina, 1994.

[2] W. Ding, F. Wu, S. Li, "Lifting-based wavelet transform with directionally spatial prediction," Picture Coding Symposium, San Francisco, CA, 2004.

[3] W. Ding, F. Wu, X. Wu, S. Li, H. Li, "Adaptive directional lifting-based wavelet transform for image coding", IEEE Trans. on Image Processing (accepted).

[4] B. Zeng and J. Fu, "Directional Discrete Cosine Transforms for Image Coding". ICME 2006.

[5] C. Loeffler, A. Lightenberg, and G. Moschytz, "Practical fast 1-D DCT algorithms with 11 multiplications", in Proc. IEEE ICASSP, vol. 2, pp988-991, 1989.

[6] J. Liang, T. D. Tran, "Fast multiplierless approximations of the DCT with the lifting scheme", IEEE trans. on Signal Processing, vol. 49, no. 12, pp 3032-3044, 2001.

[7] B. Feng, J. Xu, F. Wu, S. Yang, S. Li, "Energy distributed update steps (EDU) in lifting based Motion compensated video coding". International conference on image processing, Singapore. Oct, 2004.

[8] A. Robert, I. Amonou and B. Pesquet-Popescu, "Improving DCT-based coders through Block Oriented Transforms", 8th International Conference on Advanced Concepts for Intelligent Vision Systems, Antwerp, Belgium, Sept, 2006.

[9] H. Xu, J. Xu, F. Wu, "Lifting-based Directional DCT-like Transform for Image Coding", IEEE Trans. on CSVT (accepted).